

BEAMLINE

X19A

PUBLICATION

Beauchemin, S., D. Hesterberg, and M. Beauchemin, "Principal component analysis approach for modeling sulfur K-XANES spectra of humic acids," *Soil Sci. Soc. Am. J.*, **66**, 83-91 (2002).

FUNDING

US National Science Foundation, North Carolina Agricultural Research Service (NC-ARS)

FOR MORE INFORMATION

Suzanne Beauchemin, Research Scientist, Soil and Water Chemistry, Natural Resources Canada, Environmental Lab., Ottawa, Ontario, Canada.

Email: sbeauche@NRCan.gc.ca

Homepage: <http://www.nrcan.gc.ca/mms/canmet-mtb/mmsl.htm>

Modeling Sulfur K-XANES Spectra of Humic Acids With the Principal Component Analysis Approach

S. Beauchemin, D. Hesterberg, and M. Beauchemin

Natural Resources Canada, Ontario, Canada; North Carolina State University, Raleigh

Characterization of the chemical species of trace elements in soils, sediments, and other geochemical systems is essential for predicting their solubility and potential mobility in the environment. To determine metal speciation, scientists usually model x-ray absorption near edge structure (XANES) spectra by using standards of well-defined chemical species. But when scientists deal with complex matrices such as soil – in which many different chemical forms of the element may co-exist – other modeling techniques called principal component analysis (PCA) and target analysis, might be more convenient to use. The PCA was successfully applied to sulfur K-XANES spectra of six humic acid samples.

X-ray absorption near edge structure (XANES) spectra from unknown soil samples are usually modeled using techniques called spectral deconvolution and/or least-squares linear combination fitting (LCF). Both approaches rely on *a priori* information. Spectral deconvolution is based on an assumption about the shape of the band, while least-squares LCF involves fitting pure standard species to resolve an unknown mixture.

We tested a different modeling technique, called principal component analysis (PCA), to analyze soil mixtures. This technique considers the statistical variance within an experimental data set composed of a group of unknown samples. The data set is first redefined in terms of a reduced number of independent sources of variability, called components. These components have no chemical or physical meaning.

A subsequent analysis, the target transformation, is used to test if

suspected known species are part of the structural variation of the sample set. Based on the results from target analysis, abstract matrices obtained in the first step are transformed into chemically meaningful matrices.

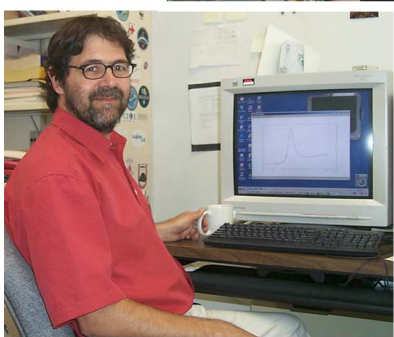
A major advantage of the PCA approach is that no *a priori* assumption is needed regarding the shape

of the band. Also, suspected species can be evaluated individually without *a priori* knowledge of other species present in the sample.

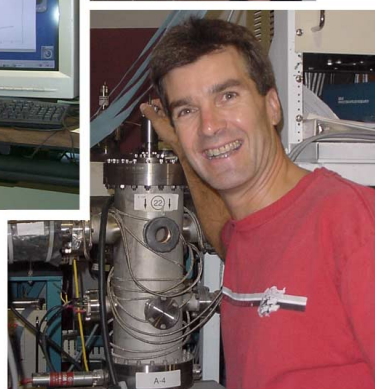
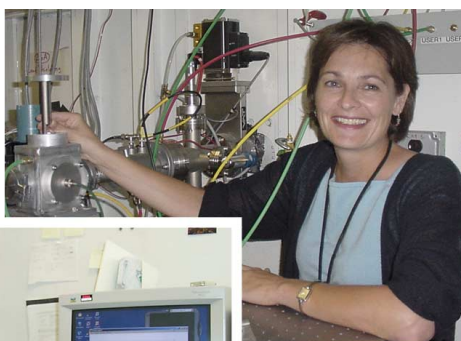
PCA coupled with target transformation was used to model sulfur K-XANES spectra of six soil humic acid samples (Figure 1) using the spectra of seven chemical compounds as sulfur standards: sodium

sulfate, chitin sulfate, cysteic acid, benzyl sulfide, benzyl disulfide, elemental sulfur, and methionine. The objective was to compare the results from PCA approach to those obtained from least-squares LCF.

PCA identified three main components. As illustrated in Figure 2, target analysis showed how closely the predicted targets fitted the suspected known species. Chitin sulfate was a better sulfate species than sodium sulfate to explain our experimental data. Sodium sulfate, cysteic acid and benzyl disulfide were marginal candidates, which means that the nature of the real species



(Authors from top, clockwise): S. Beauchemin, D. Hesterberg, and M. Beauchemin



in the spectra of humic acids may be slightly different. Other standards were rejected. Therefore, we retained chitin sulfate, cysteic acid, and benzyl disulfide to derive the relative proportion of each selected standard species in each sample.

The PCA approach provides a statistical basis for (i) gaining insight into the identification of chemical species in a multi-component sample, and (ii) defining the number of individual species to include in the spectral fitting. Therefore, it

is a valuable complementary tool for band resolution or LCF as it helps to avoid an excessive number of bands or standards in the fit.



Figure 1. Soil humic acids were extracted from a coastal salt marsh in eastern North Carolina. Soils in salt marshes such as this one near Beaufort, NC, contain reduced organic sulfide that has a high binding affinity for mercury and perhaps other trace elements. (Photo courtesy of Stephen W. Broome, NC State.)

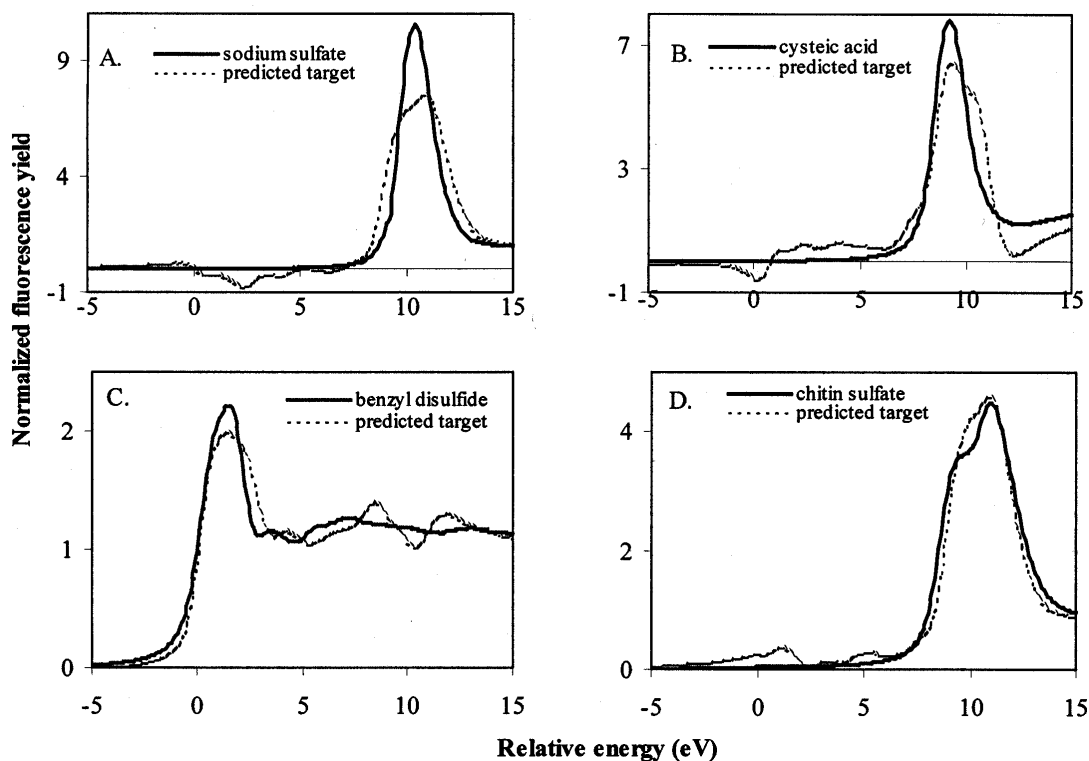


Figure 2. Predicted targets obtained through target transformation of the sulfur K-XANES spectra for (A) sodium sulfate, (B) cysteic acid, (C) benzyl disulfide, and (D) chitin sulfate compared with the normalized K-XANES spectra for these species. The energy scale is relative to S(0) K-edge at 2472 eV.